Targeted Musa Genome Sequencing and Frame Map Construction (SP2)

Collaborators: F. Bakry (CIRAD), F.C. Baurens (CIRAD), J. Dolezel (IEB), P. Heslop–Harrison (University of Leicester), I. Hippolyte (CIRAD), T. Matsumoto (NIAS), R.N.G Miller (UCB), M. Rouard (Bioversity International), N. Roux (Bioversity International), S. Sidibe-Bocs (CIRAD), M. Souza (EMBRAPA), **T. Sasaki (NIAS)**

Further information can be found online on the Global Musa Genomics Consortium website : http://www.musagenomics.org

Developing basic genomic tools to assist germplasm exploitation is important for *Musa* (banana), especially in the context of the use of *Musa* genomic diversity. Moreover, it is also a top priority in the vision of the whole genome sequencing. The initiation of genomic sequencing to identify genes will allow exploitation of breeding-relevant genetic variation within *Musa*. The GCP has promoted the development of publicly available, well characterized genomic resources in order to integrate them with ongoing trait-based research in *Musa*.

The Project

Bananas (*Musa* spp.) are **one of the developing world's most important food crops** and are grown in more than 100 countries throughout the tropics and sub-tropics.

Productivity is declining in many regions. Moreover the production systems that depend on agrochemicals and irrigation are not sustainable.

Cultivars are mostly sterile and parthenocarpic triploids (2n=3x=33), originating from two wild diploid species: *Musa acuminata* (A genome) and *Musa balbisiana* (B genome).

Despite the importance of this crop, basic genetic and genomic tools are lacking.

Objectives:

- Consolidate and characterize existing publicly available *Musa* genomic resources and make new resources available
- Selective sequencing of BACs (Bacterial Artificial Chromosome) bearing genes potentially involved in important traits
- Establishment of international mapping populations with a view to precision genetic mapping
- Anchor BACs, maps, SSR and others markers with reference to gene and genome sequences from rice and other models
- Establishment of public databases (http://www.musagenomics.org); distribution of clones; training and dissemination of information

Genetic Map

The **BORLI population** (*M.a.* 'Borneo' x *M.a.* 'Pisang Lilin') was chosen as the mapping population mainly because of the great heterozygosity of both parents (76% each) and the population size.

The map is based on 180 individuals. SSR markers were analysed on all the population, DArT markers on half on the population. 180 SSR and 380 DArT markers are distributed on 11 main linkage groups for Borneo and 9 linkage groups for 'Pisang lilin'. A consensus map was drawn for 6 linkage groups, for the other ones, translocation breakpoints and/or chromosome rearrangements did not allow map reconstruction. Analysis on these groups will be refined.

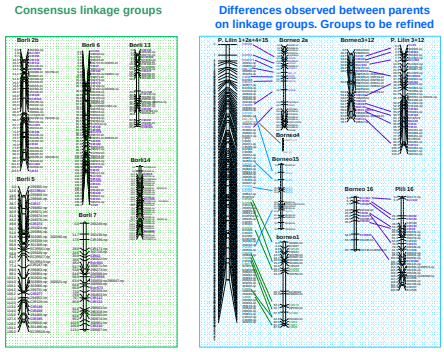

Fig. 2: Borli genetic map built from an F1 progeny of Musa acuminata 'Borneo' x Musa acuminata 'Pisang lilin'. Join map4 software was used at LOD 6 with Kosambi function. Linkage groups numbered according to Vilarhinos *et al.*, 2006

Genomic Resources and Sequencing

- 144x384 well plates of an A genome (MA4) and 96x384 well plates of the B genome (MBP PKW) BAC library have been pooled.
- 41 BACs bearing genes mainly related to abiotic and biotic stresses have been sequenced, a total of 3.303Mbp. ESTs have been used as probes to identify the BACs to sequence (cf. figure below).
- 2 cDNA libraries from PKW leaves and roots were produced. 5760 ESTs clones from these libraries have been sequenced.
- Additional 6132 ESTs from a previously developed cDNA library from *Mycosphaerella fijiensis* infected leaves of 'Calcutta 4' were produced.
- 614 repetitive DNA sequences, isolated from 'Calcutta 4', were analyzed.
- The A and B genomes are being compared, using abiotic stress related and RGA (Resistant Gene Analogue) loci to identify BACs for sequencing.

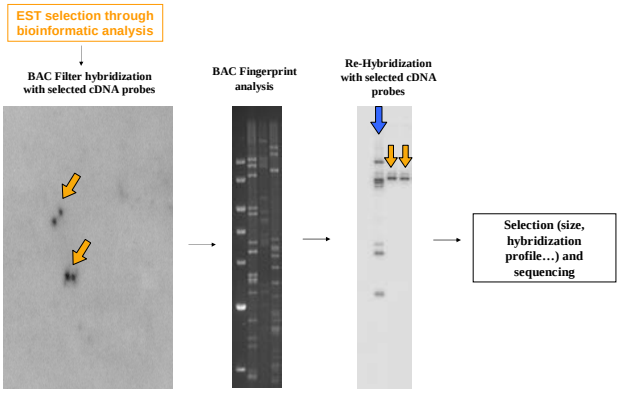

Fig. 1: Defining BACs for sequencing through EST hybridization approach

Comparative Genomics

Several annotation studies are ongoing, especially on biotic and abiotic stress. BACs showing micro-synteny (based on BAC-ends) with rice are being analysed. The most advanced study is related to RGAs and is illustrated below:

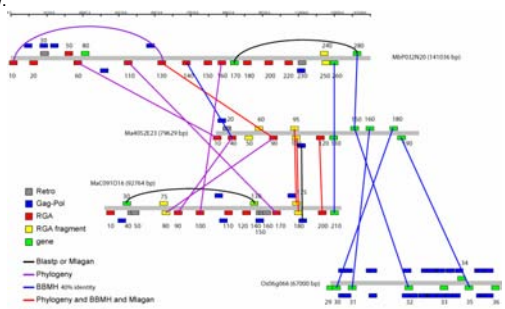

Fig. 3: Comparison of homologous regions containing clusters of duplicated RGAs within *Musa* species and with rice. Micro-synteny appears between Ma4052E23 and a region of the rice chromosome 6

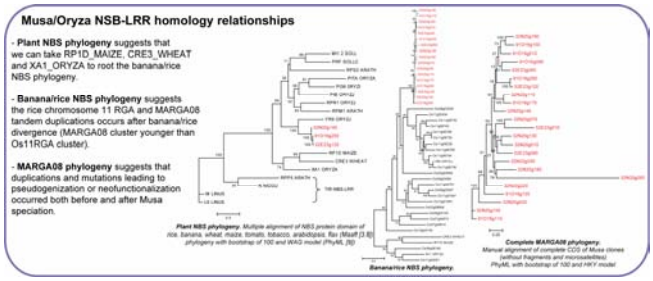

Fig. 4: *Musa/Oryza* NSB-LRR homology relationships. Banana/rice NBS phylogeny suggests that the rice chromosome 11 RGA and MARGA08 tandem duplications occured after banana/rice divergence (MARGA08 cluster younger than Os11RGA cluster)